Enhancing Communication in Noisy Environments

Craig Burrell, Ph.D.
Defence R&D Canada – Toronto
1133 Sheppard Ave West, P.O. Box 2000, Toronto, ON, Canada, M3K 2C9

Craig.Burrell@drdc-rddc.gc.ca

Sharon Abel, Ph.D.
Defence R&D Canada – Toronto
1133 Sheppard Ave West, P.O. Box 2000, Toronto, ON, Canada, M3K 2C9

Sharon.Abel@drdc-rddc.gc.ca

ABSTRACT

Military personnel when deployed in environments characterized by high-level noise require personal hearing protection devices that may limit their auditory detection, sound localization and verbal communication capabilities. This limitation may have an impact on the successful outcome of the mission, as well as personal safety. Because of the noise and also the high incidence of hearing loss, members have to shout to be heard, and there is a high probability that commands, whether delivered face-to-face or by radio, will be misunderstood. Current hearing protection and listening devices are often incompatible with other gear and may reduce situational awareness. In consequence, personnel often dispense with hearing protection to improve operational effectiveness, resulting in hearing damage. A research study is underway that explores an alternative approach that includes the use of existing in-ear communications systems which incorporate active noise reduction combined with additional signal processing algorithms. The goal of the system is to suppress background noise while enhancing speech in order to improve face-to-face communication in noisy environments, with an expectation that such a system would permit hearing protection to be worn more consistently. The proposed system makes use of audio signals collected by an array of microphones mounted on a helmet. Integration of the system into a helmet is intended to improve compatibility with regular gear, while use of an array of microphones permits sound localization, and even steering of acoustic listening beams in specific directions, while suppressing the interference from the surrounding high-level ambient noise. We review currently available hearing protection technologies, assess their strengths and weaknesses, and motivate the need for speech enhancement technologies. We then describe the prototype system which is currently under development.

1.0 INTRODUCTION

Military personnel are exposed in the course of their work to a wide range of noise environments, including high-level machine noise and impulsive sounds from explosives and firearms. As such, it is important that they wear effective hearing protection. This safety requirement competes, however, with the desire for unobstructed communication, both radio-based and face-to-face, and with the need for the preservation and enhancement of situational awareness. Hearing protection tends to impair sound detection and source localization, both of which contribute to situational awareness, and while some hearing protection devices (HPDs) can be made compatible with military radio units, effective hearing protection generally inhibits face-to-face verbal communication.

		Form Approved
Report Documentation Page		*Form Approved* *OMB No. 0704-0188*

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE **OCT 2009**	2. REPORT TYPE **N/A**	3. DATES COVERED **-**	
4. TITLE AND SUBTITLE **Enhancing Communication in Noisy Environments**		5a. CONTRACT NUMBER	
		5b. GRANT NUMBER	
		5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)		5d. PROJECT NUMBER	
		5e. TASK NUMBER	
		5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Defence R&D Canada Toronto 1133 Sheppard Ave West, P.O. Box 2000, Toronto, ON, Canada, M3K 2C9**		8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S)	
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release, distribution unlimited**			
13. SUPPLEMENTARY NOTES **See also ADA562561. RTO-MP-HFM-181 Human Performance Enhancement for NATO Military Operations (Science, Technology and Ethics) (Amelioration des performances humaines dans les operations militaires de l'OTAN (Science, Technologie et Ethique)). RTO Human Factors and Medicine Panel (HFM) Symposium held in Sofia, Bulgaria, on 5-7 October 2009., The original document contains color images.**			

14. ABSTRACT

Military personnel when deployed in environments characterized by high-level noise require personal hearing protection devices that may limit their auditory detection, sound localization and verbal communication capabilities. This limitation may have an impact on the successful outcome of the mission, as well as personal safety. Because of the noise and also the high incidence of hearing loss, members have to shout to be heard, and there is a high probability that commands, whether delivered face-to-face or by radio, will be misunderstood. Current hearing protection and listening devices are often incompatible with other gear and may reduce situational awareness. In consequence, personnel often dispense with hearing protection to improve operational effectiveness, resulting in hearing damage. A research study is underway that explores an alternative approach that includes the use of existing in-ear communications systems which incorporate active noise reduction combined with additional signal processing algorithms. The goal of the system is to suppress background noise while enhancing speech in order to improve face-to-face communication in noisy environments, with an expectation that such a system would permit hearing protection to be worn more consistently. The proposed system makes use of audio signals collected by an array of microphones mounted on a helmet. Integration of the system into a helmet is intended to improve compatibility with regular gear, while use of an array of microphones permits sound localization, and even steering of acoustic listening beams in specific directions, while suppressing the interference from the surrounding high-level ambient noise. We review currently available hearing protection technologies, assess their strengths and weaknesses, and motivate the need for speech enhancement technologies. We then describe the prototype system which is currently under development.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE	**SAR**	**13**	
unclassified	**unclassified**	**unclassified**			

We are faced, therefore, with two objectives: to provide effective hearing protection, and to provide effective communication channels and situational awareness. With conventional HPDs, these two aims are at odds, for an improvement in the former implies a decline of the latter. In such cases, the goal must be to achieve a pragmatic balance between the two. Another possibility, however, is that advanced hearing protection might be developed which permits the conflict between these two objectives to be lessened.

Electronic pass-through hearing protection (EPHP) devices attempt to proceed in this direction. EPHP devices electronically filter environmental noise, permitting only the filtered result to pass through to the listener. It may be possible to develop intelligent audio filters that preserve communication and situational awareness while still providing protection against harmful or distracting environmental noises.

In this paper, we first provide an overview of hearing protection issues in a military context. We then review existing hearing protection technologies, from simple barrier devices to more sophisticated EPHP devices. An overview of recent efforts to overcome the shortcomings of existing technologies follows, and finally we describe a project currently underway at Defence Research & Development Canada to develop a helmet-mounted speech enhancement system to promote improved face-to-face communication in noisy military environments.

2.0 HEARING PROTECTION IN A MILITARY CONTEXT

2.1   The problem of hearing damage

Generally, exposure to continuous (8-hour) sound levels in excess of 85 dBA on a daily basis or the 8-hour energy equivalent in the case of sporadic or impulse noise will result in noise-induced hearing loss after 3-4 years [1, 2]. Initially, the hearing loss will manifest as a notch in the audiogram in the region of 4 kHz [3]. This outcome reflects the natural resonance of the ear canal at 3.8 kHz and the transfer function of the middle ear [4]. Over time, the notch will deepen and the hearing loss will spread to both higher and lower frequencies. A recently published study [5] reported that by midlife (46 yr and older) 42% of a sample of Canadian Forces (CF) military members working in land, air and maritime trades had acquired a hearing loss greater than 25 dB, the clinical criterion for diagnosis of hearing loss [6]. This outcome is consistent with data collected in US army personnel in the 1970s which showed that with 15 or more years of service, the percentage of hearing-impaired soldiers exceeded 50% [7].

Noise-induced hearing loss may be prevented by either reducing the noise at source or by the wearing of personal HPDs. Reduction of noise at source is both difficult to achieve and costly. In contrast, HPDs are readily available, effective and relatively low-cost. The Canadian Forces (CF) has had a hearing conservation program in place since 1950s. Components include noise measurement, reduction of noise at source where possible, education on the hazards of noise exposure, utilization of personal hearing protection and the regular monitoring of hearing [8]. Nonetheless, the cost of claims for noise-induced hearing loss has been steadily escalating. According to Veterans Affairs Canada, the budget for audiological services in 2006 was $41 million for 49,580 individuals. This figure does not include the cost of disability pensions which would double the total outlay [9]. The Canadian military experience is similar to that of the U.S. In a review of 70,000 audiograms of U.S. Navy and Marine Corps personnel, Bohnker et al. [10] found no evidence of an improvement due to hearing conservation initiatives. The prevalence of hearing loss increased with years of service and mean values were greater than published age corrected norms, for all ages.

2.2   Impediments to use of hearing protection

Individuals working in high-level ambients, whether in military or civilian occupational settings or leisure activities, are reluctant to wear personal hearing protection. Reasons given are discomfort, difficulty fitting the device, and decreased ability to carry out auditory tasks such as the detection and localization of warning sounds, and speech communication [11]. Degradation of situational awareness may impact the success of the mission and result in casualties during military operations. Laboratory studies have confirmed that the issues raised by CF personnel are valid.   Problems with comfort and fit relate mainly to earplugs. Although a wide range of plugs varying in materials and shape are readily available for purchase, most are sold in only one size. As well, the user must rely on instructions on the packaging with respect to method of inserting the device.   Mean real-world sound attenuation is generally significantly less than the manufacturers' specifications [12]. Speech understanding in noise does not appear to be affected in individuals with normal hearing [13]. Speech and noise are decreased proportionately and the speech-to-noise ratio (SNR) remains the same. However, in those with pre-existing hearing loss, the sound attenuation provided by the device adds to the subject's raised hearing thresholds at the speech frequencies, resulting in a decrement in speech understanding.   In contrast, sound localization will be compromised in both normal-hearing and hearing impaired listeners [14]. Right-left discrimination which depends on the central encoding of interaural differences in time-of-arrival and intensity will be preserved. Both plugs and muffs will interfere with spectral cues provided by the outer ear, resulting in decrements in the accuracy of discriminating front from rearward sound sources. Typically, plugs result in a bias in perceived location towards the back and muffs towards the front.

3.0 HEARING PROTECTION DEVICES (HPDs)

3.1   Conventional HPDs

Conventional HPDs reduce ambient sounds by the same amount regardless of their level. However, the amount varies widely across makes and models, particularly for earplugs. For earmuffs, attenuation increases from about 15 dB at 0.125 Hz to about 35 dB at 1 kHz and then remains fairly stable. If well fit, earplugs generally provide relatively more attenuation (15-25 dB) below 1 kHz but are about the same above 1 kHz for highly rated devices [12]. Low-frequency attenuation may be increased by wearing a muff and plug in combination.

3.2   Level-dependent HPDs

In contrast, the attenuation provided by level-dependent HPDs will depend on the level of the ambient. These devices incorporate either limited amplification or active noise reduction (ANR), accomplished using microphones housed in one or both ear cups [15]. In the case of limited amplification, low-level signals may be amplified by up to 10 dB until a pre-set risk criterion is reached (e.g., 82 dBA). Beyond the criterion, sound attenuation will increase by 1 dB for every 1 dB increment in sound level until the passive attenuation of the muff (e.g., 35 dB) is reached. In the case of ANR, an electronic circuit housed within the muff samples and inverts the incoming waveform and adds it out of phase to the original.   Components of the two waveforms which are out of phase will cancel, thereby reducing the overall level.   ANR is limited to frequencies below 1 kHz that often characterize industrial or military environments (e.g., aircraft cockpit).

ANR is not suitable for reduction of impulsive sounds (e.g., blast and weapon's fire), since the duration of these events is not sufficient for sampling the ambient.   For these noise events, passive level-dependent

devices, muffs or plugs are recommended. These contain a precision orifice in an acoustical duct that improves transmission of low-level sounds, with the result that speech communication is minimally reduced by less than 20 dB [16, 17]. A shock wave (e.g., weapon's discharge) in the range of 80-120 dB (depending on the manufacturer and model) creates turbulent air flow in the orifice which restricts its passage, resulting in an increase in attenuation.

3.3 Advanced communication technologies

A more recent innovation in hearing protector technology is the electronic pass-through hearing protector (EPHP). This type of device consists of a pair of conventional, level-independent earplugs or earmuffs, bilateral external microphones to pick up the ambient sound, internal speakers to present these to the ears, and an electronic processing unit which will pass and possibly amplify low-level sounds, reduce high-level continuous sounds using ANR and block impulsive sounds [18, 19]. In fact, a wide range of filtering and amplification options are available to EPHP systems. Exploration of those possibilities is an active area of research, and commercial EPHP systems have begun to appear on the market (for example, from Nacre, Sylinx, and Sensear).

4.0 ENHANCING SPEECH IN NOISY ENVIRONMENTS

To better appreciate the possible scope for EPHP devices in a military context, it may be worthwhile to consider the general features of the acoustic environment of a soldier. It is characterized by several distinctive types of audio signals, as shown in Figure 1. First, there are speech signals, including face-to-face speech, radio-mediated speech, and potentially also 'speech babble' from nearby speakers. The former two sources are target signals which we wish to preserve and enhance; the latter is generally interference. Second, there is noise originating from vehicles or machinery in the vicinity. Such sources are usually dominated by low frequencies, and the spectrum may overlap with speech signals [20, 21]. Awareness of machine noise can be an important part of situational awareness, but it is often desirable to reduce the amplitude in order to protect the soldier's hearing and to improve comprehension of speech. Third, there are impulsive sounds originating from weapon fire and explosions. Perception and localization of such sounds are a critical part of situational awareness. Finally, there are other environmental sounds, some of which may be classified as background noise, and others which may be considered relevant to situational awareness.
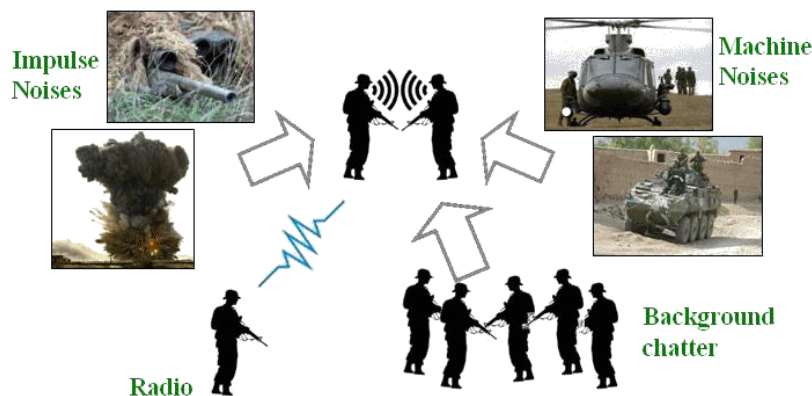

Figure 1: The acoustic environment of the soldier. Face-to-face communication competes with impulse noises, machine noises, radio communications, and other background sounds.

The hearing protection technologies discussed in Section 3 affect the soldier's perception of this acoustic environment in various ways. Conventional HPDs decrease the amplitude of all sources without discrimination. Though they are generally more effective at suppressing high frequencies [12], they are otherwise unable to make distinctions based on directionality, amplitude, source type, or other criteria. They can be effective at protecting hearing, but, for the same reason, they impair situational awareness.

Level-dependent HPDs begin to provide some discrimination between different types of sources. Limited amplification HPDs can improve situational awareness in quiet environments by amplifying environmental noises and speech, but in loud backgrounds they function in a manner similar to conventional HPDs. HPDs with ANR technology preferentially attenuate low frequency noises, which is most effective at reducing machine noise. They are not effective at identifying and improving face-to-face speech communication, nor, as mentioned above, do they provide protection against impulsive sounds.

With EPHP technology the range of possibilities is widened. Because the signals only reach the listener after passing through electronic auditory filters, the re-presentation of the auditory environment to the listener is limited only by the ingenuity of the filtering algorithms. A general aim of current research is to develop an EPHP system that provides better discrimination between target signals and interference. The system should be able to focus on signals of interest and reduce interference from competing sources. Stated in this way, the problem is a variant on the *cocktail party problem*, a challenging problem in psychoacoustics first defined by Cherry over a half-century ago [22]. Cherry noted the remarkable ability of human listeners to isolate and track a particular audio signal within a complex acoustic environment (such as the voice of an interlocutor within the complex background of voices and music at a cocktail party). The cocktail party problem is, first, the problem of understanding how the human auditory system divides the acoustic signal impinging on the ear into audio streams originating from a finite number of distinct sources, and, second, the problem of designing an automated computer system capable of performing the same task.

From a signal processing point of view, the cocktail party problem is challenging. Many attempts have been made to solve it, although by general agreement it remains an outstanding problem (see [23] for a recent review). The difficulty derives primarily from the highly non-stationary spectral characteristics of both the target and interference signals, the spectral overlapping of the target and interference, and the possible presence of reverberation, echo, and other complicating factors.

Of the various approaches that have been proposed to address the cocktail party problem, *auditory scene analysis* [24] stands out as one of the most promising, and has inspired our method. In this approach, the incoming signal is segregated into streams on the basis of a set of auditory cues, such as onset time or synchronized harmonic shifts. The particular cues employed in our system are discussed below in more detail. Once the cue-derived audio streams are established, it is possible to amplify the stream of interest and attenuate the others.

4.1   Fuzzy Cocktail Party Processor

Defence Research & Development Canada has recently initiated a project with the aim of developing an EPHP system that provides hearing protection and radio communication while enhancing face-to-face speech communication and impulsive source localization in noisy military environments. In this section we describe the design of this system, discuss the basic technical approach, and summarize early performance indicators.

The system has two independent components: a speech enhancement unit, and an impulse localization unit. Both units make use of helmet-mounted directional microphone arrays, and both are designed to comply with

a few basic requirements. The system must be wearable, and so make limited computational and power demands. The signal processing must also be carried out in real-time, which is a significant constraint. Finally, the system must be adaptive in order to have robust performance under complex, dynamic acoustic conditions. We describe the two major components of the system separately.

4.1.1   Speech enhancement unit

The speech enhancement unit has four main components: a microphone array to collect the ambient audio signal, a signal processing system to filter the signal and enhance speech, an Active Noise Reduction (ANR) component, and a hearing-protective earpiece to deliver the processed signal to the wearer. The microphone array has four directional microphones located in pairs near the ears, each pair consisting of one forward-facing and one rear-facing microphone. The speech enhancement and ANR components work co-operatively: as the SNR increases past the point where the speech enhancement unit performs effectively, it is gradually replaced by the ANR system. The hearing protective earpiece will be interfaced also with the soldier's radio communication unit.

The signal processing system, called the Fuzzy Cocktail Party Processor (FCPP), is the main innovative component of the system. It has been developed primarily by Karl Wiklund at McMaster University [25]. Its basic architecture is shown in Figure 2. The input to the system is the four-channel digital signal obtained from the directional microphones. The central processing blocks are book-ended by cochlear filterbanks which produce the frequency-domain representation of the signal prior to processing and also reconstruct the time-domain signal to be delivered to the listener. A cochlear filterbank, which consists of a set of bandpass gamma-tone filters [26], mimics the frequency decomposition performed by the human ear, and can be efficiently implemented [27].
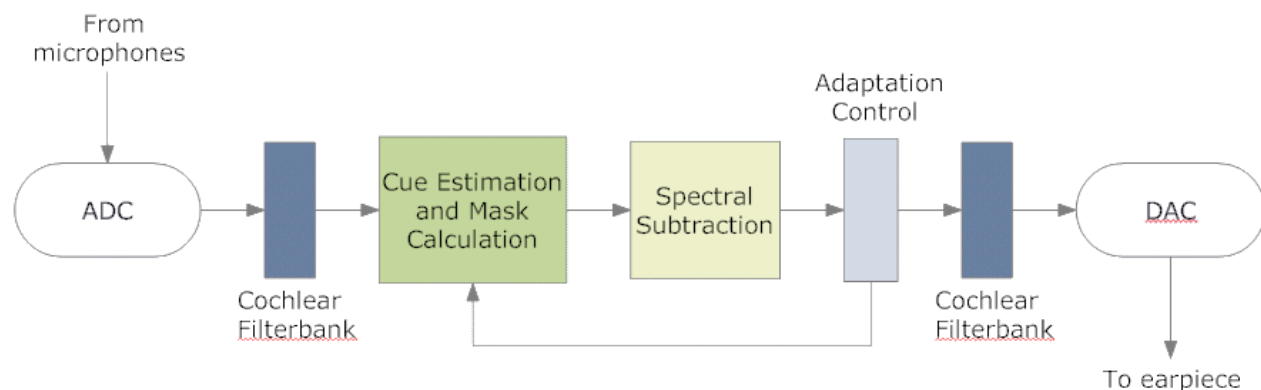

Figure 2: Fuzzy Cocktail Party Processor (FCPP) [25].

The next block, which performs cue estimation and mask calculations, is the heart of the system. In the spirit of auditory scene analysis, a set of auditory cues are used to assess the probability that a given time-frequency component of the signal belongs to the target (speech) signal. This probability is then applied as a mask on the time-frequency plane to enhance the target signal relative to the background.

The auditory cues used by the system are onset, pitch, interaural time-of-arrival difference (ITD), and interaural level difference (ILD). Onset refers to the time at which a new sound is introduced into the environment. Frequency components with correlated onsets are likely to originate from the same sound

source. Because it focuses on the first appearance of a new source, onset is fairly robust against reverberation. Pitch is a cue specially associated with speech; it is known that vowels in voiced speech give rise to a periodic pulse-pattern that occurs not only in the fundamental frequency, but also across its harmonics. The presence of such a correlated periodicity across frequency bands is a good indication that they have a common origin and should be grouped together.

Both onset and pitch are monaural cues, and so do not provide directional information. Directionality is derived from the ITD and ILD cues, which are binaural. ITD depends on the azimuthal position of the source. Similarly, ILD refers to the fact that a signal will be somewhat louder at the ear nearer the source than at the farther. Because our system is intended to enhance face-to-face communication, we use directional cues to preferentially enhance sources in the forward direction. Later versions of the system could preferentially enhance sources in some arbitrary direction, perhaps directed by eye-tracking systems.

These cues are not accorded equal weight in the analysis. Due to their robustness in complex acoustic environments, the onset and pitch cues are given priority, with the other cues acting as constraints on signal source assignments.

The cues are applied to the auditory stream using a fuzzy logic system [25]. In fuzzy logic, rather than being strictly true or strictly false, assertions take on probabilistic truth values. Fuzzy reasoning rules are based on linguistic statements that capture the basic intuitive principles of the analysis. For instance, a rule might state that if *most* cues are consistent with a source directly in front of the wearer, and if the characteristics of the sound are *likely* to be associated with speech, then there is *probably* a speaker in front of the wearer. The fuzzy logic system produces a probability that a given time-frequency unit originates from the target, and this probability is applied as a mask to enhance probable targets. The fuzzy logic approach has the merits of simplicity and computational efficiency.

All of the auditory cues used in the preceding analysis are front-back symmetric; the spectral subtraction block in Figure 2 is used to distinguish sources in front of the wearer from those behind. Recall that two oppositely-oriented directional microphones are located near each ear; the signals obtained from the rear-facing microphones are subtracted from those obtained from the forward-facing microphones [21]. Finally, an adaptation control block adjusts parameters of the system in response to changes in the acoustic environment before the signal is reconverted to a time series and delivered to the listener.

4.1.2   Impulse localization unit

The impulse localization unit is also helmet-mounted but operates independently of the speech enhancement unit. It consists of eight directional microphones uniformly distributed around the perimeter of the helmet. The localization is performed by comparing the time-of-arrival of incident impulsive acoustic peaks at the individual microphones. At the present time we assume incident plane waves, which is most accurate for sources in the far-field. The system localizes in the azimuthal plane, but not in elevation or range, and no attempt is made to identify the weapon from which the impulsive sound originated. The direction of incidence computed by the prototype system is indicated to the wearer through a hand-held visual display.

4.2   Related Work

There have been numerous attempts to apply signal processing techniques to speech enhancement and noise reduction in a 'cocktail party' environment. Techniques such as Wiener filtering [28], spatial filtering [29], and Independent Components Analysis [30, 31] have shown some success, but also suffer from certain

limitations. Weiner filtering, for instance, relies on the *a priori* ability to distinguish noise from the target signal, and performs poorly when the noise background is dynamic. For a microphone array of a given size, spatial filtering is limited in the number of sources it can distinguish, and performance degrades in the presence of reverberation. Independent Components Analysis has difficulty with non-stationary noise and converges too slowly for real-time application.

Several groups have studied speech detection in open environments by attempting to identify signatures of voiced segments (vowels) and then grouping strings of such segments together into utterances [32, 33, 34]. Different groups have focused on different characteristics of voiced segments to make the identification; the analogous component of our system is our reliance on the pitch cue. It is not clear that they would perform well under noisy and reverberant conditions, and to our knowledge performance evaluations of these proposed methods are not available.

Localization of impulsive sources using wearable microphone arrays has received considerable attention from both academic and military research groups. Usually it is assumed that, as in our case, the microphone array has a fixed geometry, but at least one group has allowed for flexible arrays [34, 35] in which the microphones dynamically learn their relative orientation by listening to sounds produced by the wearer. Onset delays are used to compute the direction to sound sources in the vicinity, and no far-field assumption is made.

Defence research organizations in several NATO member nations have dedicated effort to addressing the source localization problem. At TNO (Netherlands) work has been done to mitigate the source localization problems that arise when the ears are occluded by a helmet [19, 36]. They propose using helmet-mounted microphones and a signal processing system to digitally remove the helmet's interference, thereby restoring the "open ears" condition. This method requires the prior measurement of the Head-Related Transfer Functions (HRTF) associated with a particular helmet. This is certainly an interesting approach, but to our knowledge the performance evaluation of the system has not been published.

Another related effort originates from the US Army Research Laboratory [37], in which localization is again performed using a helmet-mounted microphone array. They consider not only single-helmet localization, but also distributed, coordinated localization employing microphones on multiple soldiers and vehicles. They also draw particular attention to the challenges that arise in urban environments, notably the fact that multiple reflections from buildings can severely obscure the direction of origin of an impulsive noise.

4.3   Performance

The system described in the preceding sections is currently under development, with the first prototype scheduled for completion in spring 2010. As such, we are unable at this time to report on the final performance characteristics. However simulations of the system that take into account the FCPP algorithm, the spatial distribution of sound sources, the reverberation of the environment, and the directional gain of the microphones have been carried out by Wiklund [25]. These simulations have been produced using the R-HINT-E virtual acoustics platform [38] and are expected to give a good approximation to the actual performance of the final system. The simulations do not include the effects of the ANR system.

The performance of the system can be evaluated both objectively, by computing the signal-to-noise ratio of the input and output signals and other technical measures, and subjectively, by studying the improvement in speech intelligibility in noise afforded by the system. An improved SNR ratio is not valuable if the speech is distorted and intelligibility reduced, so both types of measure are important.

As an objective measure the signal-to-noise ratio (SNR) is evaluated. A band-averaged SNR is used, which depends on the SNRs in various frequency bands averaged over time [25]. Test scenarios were generated by combining a single forward-located target speaker with three interfering speakers originating from different azimuthal directions; the speakers each uttered one of the standardized Hearing in Noise Test (HINT) sentences [39]. The SNR value was calculated both before and after processing by the system. The results are shown in Figure 3. Simulations were carried out with two different reverberation levels, as shown, one resembling an acoustically damped room and the other a reverberant hard-walled lecture room.
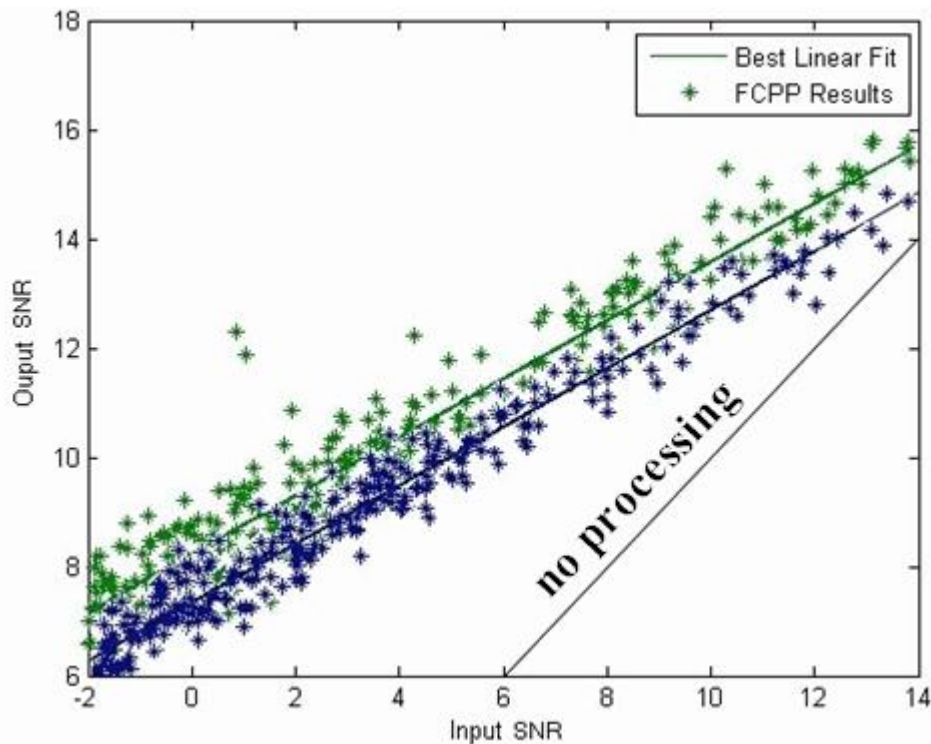

Figure 3: SNR gain from FCPP algorithm. The data in green (above) are from simulations of a low reverberation environment, and those in purple (below) from simulations of a high reverberation environment (equivalent to a hard-walled lecture room).

This simulation data showed that the system can yield good performance gain. The SNR gain decreased as the input SNR increased because the amount of noise to remove was progressively reduced. The effectiveness was somewhat better with low reverberation, as expected. The average SNR gain was 7.5 dB at high reverberation when the input SNR was 0 dB. Table 1 indicates how this result compares to that achieved with other methods. The FCPP algorithm is competitive with other methods, and is the best performing real-time method we have encountered.

Algorithm	SNR gain	Real-time processing
Multipitch tracking [40]	1.2 dB	No
Pitch tracking and amplitude modulation [41]	4.4 dB	No
Perceptual Binaural Speech Enhancement [42]	4.5 dB	Yes
Fuzzy Cocktail Party Processor [25]	7.5 dB	Yes
Binaural segregation [43]	8.9 dB	No

Table 1: SNR gains obtained with various processing algorithms. The SNR gain is given for an input SNR of 0 dB. The Fuzzy Cocktail Party Processor (FCPP) compares favourably with other published algorithms.

A pilot study was also carried out to assess the subjective benefit obtained with the FCPP algorithm [25]. Six subjects were presented with pairs of sound samples, each pair containing a sound mixture before and after processing. For each pair, the subjects were asked to compare the two samples using the Comparative Mean Opinion Score (CMOS) system [44]. A seven-level CMOS test was used, with the comparison for two signals A (processed) and B (unprocessed) ranging from "A is much worse than B" (Score=-3) to "A and B are about the same" (Score=0) to "A is much better than B" (Score=3). Each subject was presented with 80 sound mixtures. The resulting average CMOS score across subjects was 1.4 ± 0.7, indicating that the subjects considered the intelligibility of the speech in the FCPP processed signal to be "slightly better" or "better" than the unprocessed signal. This indicates that the FCPP improves the speech intelligibility to a noticeable degree.

5.0 CONCLUSION

Hearing protection is an important concern for military personnel. High noise exposures from machinery, weapon fire, and other sources are known to result in hearing damage if adequate protection is not consistently worn. Conventional hearing protection, however, also impairs situational awareness and verbal communication, with detrimental effects on operational performance and safety. An alternative approach is to use signal processing techniques to filter the ambient acoustic, attenuating background noises and enhancing signals that promote situational awareness and communication.

In this paper basic issues pertaining to hearing protection in a military context were presented, and several hearing protection technologies that are commercially available today were reviewed. A prototype system currently under development at Defence Research & Development Canada was described which makes use of Active Noise Reduction and fuzzy logic to suppress noise and enhance face-to-face verbal communication. This system is computationally efficient and performs real-time speech enhancement. Final performance data for this system will be reported when the project is completed.

REFERENCES

[1] Bauer P, Korpert K, Neuberger M, Raber A, Schwetz F. Risk factors in hearing loss at different frequencies in a population of 47, 388 noise-exposed workers. J Acoust Soc Am 1991; 90:3086-98.

[2] Pelausa EO, Abel SM, Simard J, Dempsey I. Prevention of noise-induced hearing loss in the Canadian military. J Otolaryngol 1995; 24; 271-80.

[3] Abel SM, Haythornthwaite C. The progression of noise-induced hearing loss. A survey of workers in selected industries in Canada. J Otolaryngol 1984; 13 Suppl: 2-36.

[4] Leikin JB, Davis A, Klodd DA, Thunder T, Kelafant GA et al. Selected topics related to occupational exposures. Part I. Noise exposure. Diseases of the Month 2000; 46:240-55.

[5] Abel SM. Hearing loss in military aviation and other trades: Investigation of prevalence and risk factors. Aviat Space Environ Med 2005; 76:1128-35.

[6] Yantis RA. Pure-tone air-conduction testing. In: Katz J, ed. Handbook of clinical audiology, 3ed. Baltimore: Williams & Wilkins; 1985:153-69.

[7] Henselman LW, Henderson D, Shadoan J, et al. Effects of noise exposure, race, and years of service on hearing in U.S. army soldiers. Ear Hear 1995; 16:382-91.

[8] Neely K. Hearing conservation for the armed forces. Medical Services J. – Canada 1959; April: 235-47.

[9] Veterans Affairs Canada. Personal communication. 2007.

[10] Bohnker BK, Page JC, Rovig G, Betts LS, et al. U.S. Navy and Marine Corps hearing conservation program 1995-1999: mean hearing thresholds for enlisted personnel by gender and age groups. Mil Med 2002; 167:132-5.

[11] Abel SM. Barriers to hearing conservation programs in combat arms occupations. Aviat Space Environ Med 2008; 79:591-8.

[12] Berger EH. Hearing protection devices. In Berger EH, Royster LH, Royster JD, Driscoll DP, Layne M. The noise manual, 5th ed. Fairfax, VA: Am Industr Hyg Assoc; 2000:379-454.

[13] Abel SM, Armstrong NM, Giguère C. Auditory perception with level-dependent hearing protectors: The effects of age and hearing loss. Scand Audiol 1993; 22(2):71-85.

[14] Abel SM, Hay VH. Sound localization: The interaction of aging, hearing loss and hearing protection. Scand Audiol 1996; 25(1):3-12.

[15] Abel SM, Tsang S. and Boyne S. (2007). "Sound localization with communications headsets: comparison of passive and active systems." Noise and Health 9 (37), 101-107.

[16] Allen CH, Berger EH. Development of a unique passive hearing protector with level-dependent and flat attenuation characteristics. Noise Control Eng J 1990; 34(3):97-105.

[17] SM, Lam Q. Sound attenuation of the Indoor/Outdoor Range E-A-R® plug. Mil Med 2004; 169:551-5.

[18] Brungart DS, Hobbs BW, Hamil JT. A comparison of acoustic and psychoacoustic measurements of pass-through hearing protection devices. Proceedings of 2007 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics; 2007 Oct 21-24; New Paltz, NY.

[19] Vos WK, Bronkhort AW, Verhave JA. Electronic pass-through hearing protection and directional hearing restoration integrated in a helmet. Proceedings of ASA-SFA-EAA Joint Conference, Acoustics'08; 2008 June 29-July 4; Paris, France.

[20] Winberg M. Speech enhancement and noise combat. Ph.D. thesis, Karlskrona/Ronneby, 1998.

[21] Hansler E, Schmidt G. Acoustic Echo and Noise Control: A Practical Approach. Hoboken, NY: Wiley, 2004.

[22] Cherry EC. Some experiments on the recognition of speech, with one and two ears. Journal of the Acoustic Society of America 1953, 25:319-36.

[23] Haykin S, Chen Z. The Cocktail Party Problem. Neural Computation 2005. 17:1875-1902.

[24] Bregman A. Auditory Scene Analysis: The Perceptual Organization of Sound. Cambridge, MA: MIT Press, 1990.

[25] Wiklund K. The Cocktail Party Problem: Solutions and Applications. Ph.D. thesis, McMaster University, 2009.

[26] Patterson RD, Robinson K, Holdsworth J, McKeown D, Zhang C, Allerhand MH. Complex sounds and auditory images. Auditory Physiology and Perception. Oxford, 1992, pp.429-46.

[27] Slaney M. An Efficient Implementation of the Patterson-Holdsworth Auditory Filter Bank. Apple Computer Technical Report 35, 1993.

[28] Quateieri TF. Discrete-Time Speech Signal Processing. Upper Saddle River, NJ: Prentice-Hall, 2002.

[29] Van Trees HL. Optimum Array Processing. Hoboken, NJ: Wiley, 2002.

[30] Bell AJ, Sejnowski TJ. An information-maximization approach to Blind Source Separation. Neural Computation 1995, 16:1129-59.

[31] Haykin S. Neural Networks: A Comprehensive Foundation. Upper Saddle River, NJ: Prentice-Hall, 1999.

[32] Junqua JC, Mak B, Reaves B. A robust algorithm for word boundary detection in the presence of noise. IEEE Transactions on Speech and Audio Processing, 1994, Vol.2 No.3: 406-12.

[33] Wu GD, Lin CT. Word boundary detection with mel-scale frequency bank in noisy environment. IEEE Transactions on Speech and Audio Processing, 2000: Vol.8 No.5: 541-53.

[34] Basu S, Clarkson B, Pentland A. Smart Headphones: Enhancing Auditory Awareness through Robust Speech Detection and Source Localization. Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2001: 3361-4.

[35] Basu S, Schwartz S, Pentland A. Wearable Phased Arrays for Sound Localization and Enhancement. Proceedings of the 4th IEEE International Symposium on Wearable Computers, 2000: 103-10.

[36] Bronkhorst AW, Verhave JA. A microphone-array-based system for restoring sound localization with occluded ears. New Directions for Improving Audio Effectiveness: 20-1 – 20-8.

[37] Scanlon MV. Helmet-mounted acoustic array for hostile fire detection and localization in an urban environment. Proceedings of SPIE, 2008: Vol.6963, D-1 – D-13.

[38] Wiklund K, Sonnadara R, Trainor L, Haykin S. R-HINT-E: A realistic hearing in noise test environment. ICASSP 2004, May 17-21, 2004. Montreal, QC.

[39] Nilsson M, Soli SD, Sullivan JA. Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. J. Acoust. Soc. Am. 1994, 95:1085-99.

[40] Wu M, Wang D, Brown GJ. A multipitch tracking algorithm for noisy speech. IEEE Transactions on Speech and Audio Processing. 11 (2003): 229-36.

[41] Wu G, Wang D. Monaural speech segregation based on pitch tracking and amplitude modulation. IEEE Transactions on Neural Networks. 15 (2004): 1135-50.

[42] Dong R. Perceptual Binaural Speech Enhancement in Noisy Environments. M.A.Sc. Thesis, McMaster University, 2004.

[43] Roman N, Srinivasan S, Wang D. Binaural segregation in multisource reverberant environments. Journal of the Acoustical Society of America. 120 (2006): 4040-51.

[44] Dreiseitel P, Schmidt G. "Evaluation of Algorithms for Speech Enhancement", Jens Blauert, Ed. Communication Acoustics. Berlin: Springer-Verlag, 2005.